

論文内容の要旨

博士論文題目 Context Enhancement of Recurrent Neural Network Language Models for Automatic Speech Recognition
(音声認識用ニューラルネットワーク言語モデルのコンテキスト強化)

氏名 Michael Alexander Hentschel

(論文内容の要旨)

言語モデルは、自動音声認識システムの重要な要素である。近年、ニューラルネットワーク、特にリカレントニューラルネットワークを用いている言語モデルは、従来のカウントベース言語モデルの性能を大幅に改善した。言語モデルは、単語履歴から次の単語の確率を計算するモデルである。自動音声認識のタスクでは、単語以外に様々なコンテキスト情報が利用可能である。このコンテキストは、音響信号、最近の単語履歴、またはグローバルテキストトピックの様な異なる情報源に基づくことができる。しかし、ニューラルネットワーク言語モデルは、この種のコンテキスト情報を特に利用していない。本論文では、自動音声認識において言語モデルが利用可能な種類のコンテキスト情報を提案する。これらの情報を言語モデルで利用可能にするため、コンテキスト情報を利用するネットワークアーキテクチャを提案する。続いて、最新のリカレントニューラルネットワーク言語モデルへの有効性に対して、コンテキスト情報とネットワークアーキテクチャとの様々な組み合わせを調査する。提示された手法は、モデルのパープレキシティとリスコアリングでのワードエラーレートで最新のリカレントニューラルネットワーク言語モデルの改善を達成した。

氏名	Michael Alexander Hentschel
----	-----------------------------

(論文審査結果の要旨)

平成30年12月27日に開催した公聴会の結果を参考に平成31年2月13日に本博士論文の審査を行った。以下のとおり、本博士論文は、提案者が独立した研究者として、研究活動を続けていくための十分な素養を備えていることを示すものと認める。

Michael Alexander Hentschel は、本博士論文において、自動音声認識のための言語モデルに対して様々な文脈情報を利用するニューラルネットワークモデルを提案し、その有効性を示した。

本学位論文の貢献は、次のようにまとめることができる。

1. 言語モデルは、直前の単語列を基に次の単語の生起確率を予測するが、単語列以外の様々な文脈情報も、次の単語の予測に影響を与えられられる。そのような文脈として、特に、音響信号、性別や話者同定など話者に関する情報がある。さらに、長い文脈を認識しようとした場合は、これまでの発話列があり、対話のトピックの判定に役立つと考えられる。本論文では、このような様々な文脈情報を、再帰ニューラルネットワークに基づく言語モデルの性能向上に利用する方法を提案し、有効性を確認した。
2. 具体的には、これらの文脈情報を教師なしの設定で素性として抽出する手法を提案し、様々な文脈情報を素性として利用することを可能にするニューラルネットワークアーキテクチャを提案した。
3. さらに、ニューラルネットワークを用いてこれらの文脈情報を組み合わせて言語モデルに利用する統合的アプローチを提案し、言語モデルのパープレキシティとリスクアリングでのワードエラーレートにおいてこれを評価し、これまでで最高性能を示すリカレントニューラルネットワーク言語モデルを実現した。

言語モデルの改善は、自動音声認識を始めとする音声・言語解析の基盤技術であり、広い文脈情報の利用により、その改善を実現した本研究は、独創性が高く、実用的であり、言語処理の分野において高い貢献があると評価する。

よって、本論文は、博士（工学）の学位論文として価値あるものと認める。